The Reliability of Self-Reported Cigarette Consumption in the United States

EVRIDIKI J. HATZIANDREU, MD, DRPH, JOHN P. PIERCE, PHD, MICHAEL C. FIORE, MD, MPH, VERNER GRISE, PHD, THOMAS E. NOVOTNY, MD, AND RONALD M. DAVIS, MD

Abstract: To investigate the possibility that self-reported smoking is not a valid measure for assessing trends in smoking prevalence, we compared total self-reported cigarette consumption with the adjusted consumption data from cigarette excise taxes as reported by the US Department of Agriculture (USDA) for the period 1974 through 1985. Self-reported consumption was calculated by using data from the National Health Interview Surveys (NHIS) for adults and from the National Household Surveys on Drug Abuse for adolescents. For this period, the average ratio of self-reported cigarette consumption to the USDA estimate of consumption was 0.72 (range = 0.69 to 0.78). There was no statistical difference in this consumption ratio from year to year, indicating no apparent increase in the underreporting of cigarette smoking in these surveys. We conclude that cross-sectional surveys of self-reported smoking status remain a reliable surveillance tool for monitoring changes in population smoking behavior. (*Am J Public Health* 1989; 79:1020–1023.)

Introduction

Cigarette smoking has been identified as one of the major public health hazards of the 20th century. There are two standard approaches to measure the degree of exposure of the United States population to this hazard. The first approach uses information obtained in the process of collecting state and federal excise taxes and is an aggregate measure of population consumption. These data are regularly reported as total and per capita cigarette consumption by the Economic Research Service of the US Department of Agriculture (USDA). The advantage of the aggregate measure is that it is based on objective data that would not be affected by changing social attitudes towards smoking. The second approach is to survey a representative sample of the public and to ask questions regarding individual smoking behavior. This has been undertaken on a regular basis since the mid-1960s for the population of the United States by the National Center for Health Statistics and has been the basis of trend analyses of smoking reported elsewhere.[1–3] This latter measure, however, depends on the honesty and accuracy of respondents in reporting their own behavior.

The social acceptability of a behavior has been known to affect a person's willingness to admit engaging in that behavior[4] and may also affect the accuracy in reporting the number of cigarettes smoked per day. The social acceptability of smoking has decreased with decreasing smoking prevalence and increasing restrictions on smoking in public places.[5]

Using data from four national smoking surveys undertaken between 1964 and 1975, Warner concluded that the self-reported survey estimates of cigarette consumption increasingly underestimate consumption estimates from USDA during the period 1964 through 1975.[6] While he noted that this may have resulted from a change in survey methodology, he suggested that the finding was real and resulted from the increasing social stigma and personal health risk attached to smoking. If there has been increasing underreporting of smoking behavior, then the dramatic decrease in self-reported smoking prevalence reported elsewhere may be, in part, artifactual.[1,2]

Policy decisions with regard to smoking rely heavily on data from national surveys. It is essential therefore to document the validity and reliability of this method of measurement. In this report, we compared trends in the USDA estimates of cigarette consumption from 1974 to 1985 to estimates of consumption obtained from self-reported survey data collected during the same period by the National Center for Health Statistics and the National Institute on Drug Abuse. We also examined whether the trends in per capita consumption match the decline in smoking prevalence.

Methods

To calculate cigarette consumption, data from multiple sources were used. USDA publishes estimates of aggregate cigarette consumption. To estimate self-reported consumption, smoking behavior data from the National Health Interview Surveys and the National Household Surveys on Drug Abuse were analyzed. These data sources are described below.

USDA Estimate

The Economic Research Service of the US Department of Agriculture has made estimates of total and per capita consumption of cigarettes for a number of years. The estimates are based on data from the Bureau of Alcohol, Tobacco and Firearms of the US Department of Treasury, the Bureau of Commerce of the US Department of Commerce, the Tobacco Institute, and other private industry sources.[7]

The Tobacco Institute reports the number of packs of cigarettes on which state taxes are paid; the Bureau of Alcohol, Tobacco and Firearms reports the number of cigarettes on which federal taxes are paid; and the Bureau of Commerce reports the number of cigarettes imported into the United States. Both federal and state excise taxes are collected at the wholesale level (on removal from warehouses) and are not sales taxes. The tax-based data are then adjusted for estimated inventory changes and for returned unused products. Adult per capita consumption is calculated by dividing total consumption by the total US population 18 years of age and older for any given year.

Self-Reported Adult Consumption

The National Center for Health Statistics (NCHS), through the National Health Interview Survey (NHIS), has

From the Office on Smoking and Health, Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control. Address reprint requests to John P. Pierce, PhD, Chief, Epidemiology Branch, Office on Smoking and Health, CCDPHP, CDC, 5600 Fishers Lane, Rockville, MD 20857. Dr. Grise is with the US Department of Agriculture; Dr. Fiore is with the Department of Medicine, University of Wisconsin, Madison. This paper, submitted to the Journal October 18, 1988, was revised and accepted for publication February 7, 1989.

collected and made publicly available health information from a probability sample of the civilian, noninstitutionalized population of the United States since 1964. In cross-sectional surveys in 1974, 1976, 1978, 1979, 1980, and 1985, NCHS administered a supplemental questionnaire on tobacco use to a subsample of the population surveyed. These surveys included adults aged 18 years and older in 1983 and 1985, aged 19 years and older in 1976, and aged 17 years and older in 1974, 1978, 1979, and 1980. Detailed descriptions of the NHIS design and procedures have been published.[8,9] While the sample design of the NHIS changed in 1985, the changes were aimed at making the population estimates more precise. They do not invalidate comparisons among the estimates for the different survey years.

These supplemental surveys on tobacco use include self-respondents only (proxy responses were not allowed). When a selected respondent could not be interviewed in person, the survey was administered by telephone. The average response rate for the tobacco use supplements was 89.8 percent, with a range of 88.5 percent to 91.1 percent (NCHS unpublished data) and the sample size ranged from 9,553 in 1980 to 31,082 in 1985.[1]

The data were classified into age-sex-race groups and weighted to reflect the distribution of the US population for the survey year and the individual probability of selection as well as other survey design features. Ever smokers were defined as persons who reported that they had smoked at least 100 cigarettes in their lifetime. They were further identified as current smokers or former smokers based on responses to the question "Do you smoke now?" Respondents were asked how many cigarettes they smoked per day. We included in the analysis only survey data collected after 1974 because data collected prior to 1974 included proxy responses. Proxy reporting has been shown to underestimate the number of cigarettes smoked per day.[10]

Self-Reported Adolescent Consumption

Data on adolescents' smoking behavior are not available from the NCHS. However, the National Institute on Drug Abuse (NIDA) has conducted national household surveys of the civilian, noninstitutionalized population aged 12 years and older in the contiguous United States (Alaska and Hawaii excluded) in 1974, 1976, 1977, 1979, 1982 and 1985. These surveys use a self-administered questionnaire, administered in the household, and are based on a national probability sample of households. Respondents were asked the average number of cigarettes smoked during the past 30 days. The data are weighted to reflect selection probabilities based on age, sex, and race, and adjustments are made for nonresponses. The sample size in these surveys for those aged 12-17 years was approximately 2,000 for each survey year.

Analysis

Consumption Estimates—For each year, the total cigarette consumption is the amount smoked by: adult current smokers, adolescent current smokers, and former smokers who stopped smoking during that year. For each of these groups, the daily consumption is the product of the following three variables: 1) proportion of the group that is smoking (smoking prevalence); 2) the population size in each group; and 3) the average number of cigarettes smoked per day by each smoker in the group. Therefore, in each survey year, the consumption for each group equals the reported daily consumption times 365 days per year.

Calculations for adults were made using the public use data tapes available from NCHS for the years in which the supplements on smoking were conducted. Data for the adolescents were estimated from the NIDA surveys. We relied on published data for all years except 1985[11]; unpublished 1985 data were provided by NIDA (Joe Gfroerer, personal communication). Because the NIDA survey years differ from those of the NHIS, the prevalence and the average cigarette consumption were estimated by using linear intrapolation to match the NHIS years.

Census data were used to provide estimates of the appropriate age-specific population in the United States for each survey year.[12]

For current smokers covered by the NHIS, calculating the yearly consumption estimates was straightforward, using the smoking prevalence and the reported number of cigarettes smoked per day. Consumption by former smokers who quit in the year preceding the survey was estimated by using data from the combined 1978–1980 NHIS. For these three years, former smokers who quit during the previous year smoked an average of 192 days during that year, very close to half of the year. Accordingly, we assumed that former smokers who quit prior to the survey smoked for half of the year.

Because some adolescents also began to smoke during the year of the survey, self-reported cigarette consumption during the 30 days prior to the survey multiplied by 12 months overestimates adolescent yearly consumption. We had no valid and reliable national data to adjust for this overestimate. On the other hand, our data underestimate actual adolescent consumption because they do not include consumption by "experimental smokers" who had smoked during the year but had not done so in the 30 days prior to the survey. We conducted a sensitivity analysis to determine how assumptions made in calculating cigarette consumption affected the results of the analysis.

Consumption Ratios—Using both sources of cigarette consumption, we calculated a consumption ratio for each year by dividing the total self-reported cigarette consumption by the USDA estimate of consumption. This measure expresses the self-reported data as a fraction of the USDA estimate.

Statistical Analysis—Data were analyzed using STATISTIX,[13] an interactive program for microcomputers. A least squares line of best fit was calculated for both the per capita consumption and for the consumption ratios.

Results

The USDA estimate of cigarette consumption in the United States increased steadily from 1974 through 1981, when an estimated total of 640 billion cigarettes were smoked. Since 1981, this estimate of cigarette consumption has steadily declined despite increasing population size; in 1986, the number of cigarettes consumed was 584 billion (Table 1).

Per capita consumption decreased steadily from 4,141 cigarettes per year in 1974 to 3,274 in 1986 (Table 1). A regression line was fitted to these data and this line described the data well ($R^2 = 0.95$). The average reduction in per capita consumption was 75.4 cigarettes per year, (95% CI = 65.1, 85.7).

The USDA estimate of total cigarette consumption, self-reported cigarette consumption (based on NCHS and NIDA surveys), and the consumption ratios for the selected years from 1974 through 1985 are shown in Table 2. Although self-reported consumption is less than the USDA estimate of

TABLE 1—USDA Estimates of Total Cigarette Consumption and per Capita Consumption, United States, 1974 through 1986

Year	Estimated Total Cigarettes Consumed per Year (billions)	Estimated* per Capita Consumption of Cigarettes Per Year
1974	599.0	4,141
1975	607.2	4,123
1976	613.5	4,092
1977	617.0	4,051
1978	616.0	3,967
1979	621.5	3,861
1980	631.5	3,844
1981	640.0	3,836
1982	634.0	3,739
1983	600.0	3,488
1984	600.4	3,446
1985	594.0	3,370
1986	583.8	3,274

*Based on US population 18 years of age and older; manufactured cigarettes only.
SOURCE: US Department of Agriculture.

consumption at each year, this relationship appears to be constant over time. The consumption ratios ranged from 0.78 in 1983 to 0.69 in 1976, with a mean of 0.72. The slope of regression line is not significantly different from zero ($p = 0.85$), strongly suggesting the lack of any change over time in the consumption ratio. The components of the self-reported cigarette consumption data for each year and for each of the three groups are shown in Table 3. The contributions of adult former smokers and adolescent smokers are small relative to the amount smoked by adult current smokers.

Sensitivity Analysis

For the former-smoker and the adolescent components of total self-reported consumption, assumptions that were used in calculating self-reported consumption were:

- former smokers who quit during the preceding 12 months quit on average at the mid-point of the year; therefore, they were smoking for 182 days (rather than 192 days estimated from the 1978–80 surveys);
- adolescent data could be interpolated from the NIDA survey years to the NHIS years;
- the impact of adolescents who took up smoking during the preceding year and the impact of the experimental smokers who did not smoke in the 30 days prior to the survey were small and canceled each other out.

The results were not altered when 192 days was used instead of 182 days. The mean consumption ratio was

TABLE 2—USDA and Self-Reported Cigarette Consumption Estimates, United States, Selected Years 1974 through 1985

Year	USDA Estimate (billions)	Estimate from Self-Reported Data (billions)	Consumption Ratio
1974	599.0	434.86	0.726
1976	613.5	424.36	0.692
1978	616.0	438.38	0.712
1979	621.5	441.20	0.710
1980	631.5	459.11	0.727
1983	600.0	467.83	0.780
1985	594.0	414.35	0.698

SOURCES: US Department of Agriculture, National Center for Health Statistics (National Health Interview Survey), National Institute on Drug Abuse (National Household Survey on Drug Abuse).

TABLE 3—Self-Reported Cigarette Consumption among Different Population Sub-groups, United States, Selected Years, 1974 through 1985

	Estimated Cigarette Consumption in Billions by			
Year	Adult Current Smokers	Adult Former Smokers*	Adolescent Current Smokers	Total Population
1974	393.97	21.38	19.51	434.86
1976	387.34	20.30	16.72	424.36
1978	404.02	19.47	14.89	438.38
1979	405.59	22.44	13.17	441.20
1980	423.80	23.02	12.29	459.11
1983	432.43	25.02	10.38	467.83
1985	377.82	26.26	10.26	414.35

*Former smokers who had quit smoking during the preceding year.
SOURCES: National Center for Health Statistics (National Health Interview Survey), National Institute on Drug Abuse (National Household Survey on Drug Abuse).

affected at the third decimal place (0.720 versus 0.721). This lack of effect is partially due to the small contribution of former smokers to the total amount of the self-reported consumption. The contribution of the adolescent smoking component to total consumption was even smaller than that of former smokers for each survey year (Table 3). It ranges from 2.2 percent to 4.5 percent of the total self-reported consumption. If it would increase by 50 percent at each year, the mean consumption ratio would be affected at the second decimal place (0.720 vs 0.732); therefore variations in this component are not likely to influence the results substantially either. Given this small impact of adolescent smoking, an overlap between the two surveys for the lower age bound (those aged 17)—which occurred in the years 1974, 1978, 1979, and 1980—is not likely to have a major effect. Overall, over 90 percent of the total self-reported consumption is attributed to current smokers for whom relevant data are readily available.

Discussion

Measurement of any quantity raises issues of validity and reliability. The findings of this analysis, comparing USDA estimates of cigarette consumption to self-reported smoking consumption, 1974 through 1985, address both questions. The consumption ratio 0.72 represents a considerable discrepancy between the two measures. It is unlikely that this discrepancy between the USDA and the self-report estimates stems from the former and adolescent components of the total consumption because they account for less than 10 percent of the total consumption. It may be that adult current smokers do not report their smoking status correctly, but the literature does not support this conclusion. Findings from biochemical validation studies[14,15] suggest that self-reported cigarette consumption prevalence is a valid estimate of smoking status in the population. An alternative explanation may be that current smokers consistently underreport the number of cigarettes they smoke. This underreporting may occur disproportionately among those who are heavier smokers, analogous to the underreporting of alcohol intake.[16]

Alternatively, it might be from a systematic rounding down of the number of cigarettes smoked per day as respondents have a preference to report daily consumption in multiples of a half-pack of cigarettes and therefore report the closest half-pack consumption below their actual consumption level.[5,17]

The ratio between the two estimates is constant over time, strongly suggesting that self-reported data from these national surveys are reliable, despite any increasing social unacceptability of smoking which could theoretically bias the accuracy of self-reported data. A bias could occur if the two components of the self-reported consumption estimate, i.e., prevalence and daily consumption, change differentially over time. Prevalence has been decreasing steadily since 1974.[1,2] Daily consumption, however, remained relatively constant over time.[5]

These findings differ from the conclusions drawn from the study by Warner[6] who reported that the consumption ratio decreased from 0.73 in 1964 to 0.64 in 1975. This earlier analysis used data from the 1964, 1966, 1970, and 1975 Adult Use of Tobacco Surveys. These surveys, however, used different methodologies, particularly with respect to the mode of data collection. The 1964 and 1966 surveys used in-person interviews, whereas the 1970 and 1975 surveys used telephone. The consumption ratios obtained for 1964 and 1966 (0.73 and 0.72, respectively), using in-person survey data, are similar to the mean ratio of 0.72 observed in our study from in-person surveys during the period 1974 to 1985. Therefore, there is no evidence for an increase in underreporting when comparing in-person survey data from the 1960s with data from 1974 to 1985. The drop in consumption ratio reported by Warner between 1966 and 1970 is probably best explained by the change in the interview mode, because telephone surveys generally provide slightly lower estimates than in-person household surveys.[5]

One might ask whether it is more appropriate to use the USDA estimates of cigarette consumption, which are more objective and easily available, or self-reported data such as the NCHS or the NIDA surveys, as the primary source for surveillance of smoking in the population. The USDA estimate is a single overall figure. However, such a figure does not provide the number and types of people who smoke or the number of cigarettes each smoker consumes and hence cannot be used to target antismoking programs. On the other hand, trend analyses of self-reported smoking[1–3] do provide the smoking behavior of individuals in the population (and in sociodemographic subgroups), both in terms of smoking status and daily consumption patterns. The aggregate estimate may be more sensitive to influences of anti-smoking activity than estimates of smoking prevalence because, for many people, it may be easier to reduce the number of cigarettes smoked per day than to quit smoking. This aggregate estimate, however, cannot distinguish between a large number of people reducing the amount they smoke and a smaller number who actually quit.

The most important public health goal is to reduce smoking prevalence, not simply to reduce the number of cigarettes consumed by each smoker. Even though a reduction in daily cigarette consumption has some health benefits, these are smaller than those of quitting smoking.[5,18] Our results support the use of self-reported cigarette smoking from national surveys as a reliable surveillance tool, essential in monitoring progress toward achieving the 1990 health Objectives for the Nation,[19] and a smoke-free society by the year 2000.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the staff of the National Center for Health Statistics, and particularly Dr. Jacob Feldman, Dr. Joel Kleinman, and Ronald Wilson for collecting and providing the basic data and consulting with us on the complex design issues and methods related to the analyses; Joseph Gfroerer of the National Institute of Drug Abuse for providing the basic data; and Dr. Kantilal Patel for his statistical support.

REFERENCES

1. Fiore MC, Novotny TE, Pierce JP, Hatziandreu EJ, Patel KM, Davis RM: Trends in cigarette smoking in the United States: The changing influence of gender and race. JAMA 1989; 261:49–55.
2. Pierce JP, Fiore MC, Novotny TE, Hatziandreu EJ, Davis RM: Trends in cigarette smoking in the United States: Educational differences are increasing. JAMA 1989; 261:56–60.
3. Pierce JP, Fiore MC, Novotny TE, Hatziandreu EJ, Davis RM: Trends in cigarette smoking in the United States: Projections to the year 2000. JAMA 1989; 261:61–65.
4. Bandura A: Social Foundation of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall, 1986.
5. US Department of Health and Human Services: Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General. Public Health Service, Centers for Disease Control, Office on Smoking and Health. Washington DC: Govt Printing Office, 1989.
6. Warner KE: Possible increases in the underreporting of cigarette consumption. J Am Statist Assoc 1978; 73:314–318.
7. US Department of Agriculture: Tobacco Situation and Outlook Report. Technical Series 196. Washington DC: USDA, September 1986.
8. National Center for Health Statistics: The National Health Interview Design, 1973–84, and Procedures 1975–83. Programs and Collection Procedures Vital Health Statistics Series 1, No. 18. Public Health Service, US Department of Health and Human Services. Pub. No (PHS) 85-1320. Washington, DC: Govt Printing Office, 1985.
9. National Center for Health Statistics: Massey JT, Moore T, Todros W, Parsons V: Sample Design and Estimation Procedures for the National Health Interview Survey, 1985–1994, Vital and Health Statistics, Series 2. Public Health Service, US Department of Health and Human Services. Hyattsville, MD: NCHS, (in press).
10. Rogot E, Reid DD: The Validity of data from next-of-kin in studies of mortality among migrants. Int J Epidemiol; 4(1):51–57.
11. DHHS Pub. No. (ADM) 84-1263. National Institute on Drug Abuse: National Survey on Drug Abuse: Main Findings. Rockville MD: NIDA, 1984.
12. US Bureau of the Census: United States Population Estimates, by Age, Sex and Race: 1980 to 1987. Current Population Reports, Series P-25, No 1022. Washington, DC: Govt Printing Office, 1988.
13. NH Analytical Software: STATISTIX II. Roseville, MN: The Company, Fall 1987.
14. Pierce JP, Aldrich PN, Hanratty S, Dwyer T, *et al:* Uptake and quitting smoking trends in Australia 1974–1984. Prev Med 1987; 16:252–260.
15. Fortmann SP, Rogers T, Vranizan K, Haskell WL, *et al:* Indirect measures of cigarette use: Expired-air carbon monoxide versus plasma thiocyanate. Prev Med 1984; 13:127–135.
16. Room MR: Measurement and distribution of drinking patterns and problems in general populations. *In:* Edwards G, Gross MM, Keller M, Moser J, Room R (eds): Alcohol Related Disabilities. Geneva: World Health Organization, 1977; 61–87.
17. Kozlowski LT: Pack size, reported smoking rates and public health. Am J Public Health 1986; 76:1337–1338.
18. DHHS: The Health Consequences of Smoking: The Changing Cigarette. A report of the Surgeon General. DHHS Pub. No. (PHS)81-50156. Rockville, MD: Office on Smoking and Health, 1981.
19. US Department of Health and Human Services: The 1990 Health Objectives for the Nation: A Midcourse Review. Washington, DC: Office of Disease Prevention and Health Promotion, Public Health Service, November 1986.